# *Predicting the effect of reference population on the accuracy of across and multi breed genomic prediction*

Irene van den Berg[1], THE Meuwissen[2], IM Macleod[3], and ME Goddard[1,3]

[1]Faculty of Veterinary & Agricultural Science, University of Melbourne, Victoria, Australia; [2]Norwegian University of Life Sciences, Ås, Norway; [3]Agriculture Victoria, AgriBio, Centre for AgriBioscience, Bundoora, Victoria 3083, Australia

# *Introduction*

- Accuracy of genomic prediction usually estimated by cross-validation $\rightarrow$ requires dataset & analyses

- It would be useful to predict the benefit of multi-breed reference populations, before having the full dataset / analysis

- Formulae to predict accuracy based on size reference population, heritability and number of independent chromosomal segments (Me)

- Difficult to estimate Me, especially in a multi/across breed setting

# *Introduction*

- Accuracy of genomic prediction usually estimated by cross-validation → requires dataset & analyses

- It would be useful to predict the benefit of multi-breed reference populations, before having the full dataset / analysis

- Formulae to predict accuracy based on size reference population, heritability and number of independent chromosomal segments (Me)

- Difficult to estimate Me, especially in a multi/across breed setting

**Objective**

*Predict the accuracy of a multi breed reference population*

# *Data*

- ~ 1 million sequence variants, simulated phenotypes

- Reference population:
  - 7,878 Holstein
  - 2,323 Jersey
  - Holstein + Jersey

- Validation populations:
  - 582 Holstein
  - 551 Jersey
  - 870 Red Holstein
  - 114 Australian Red

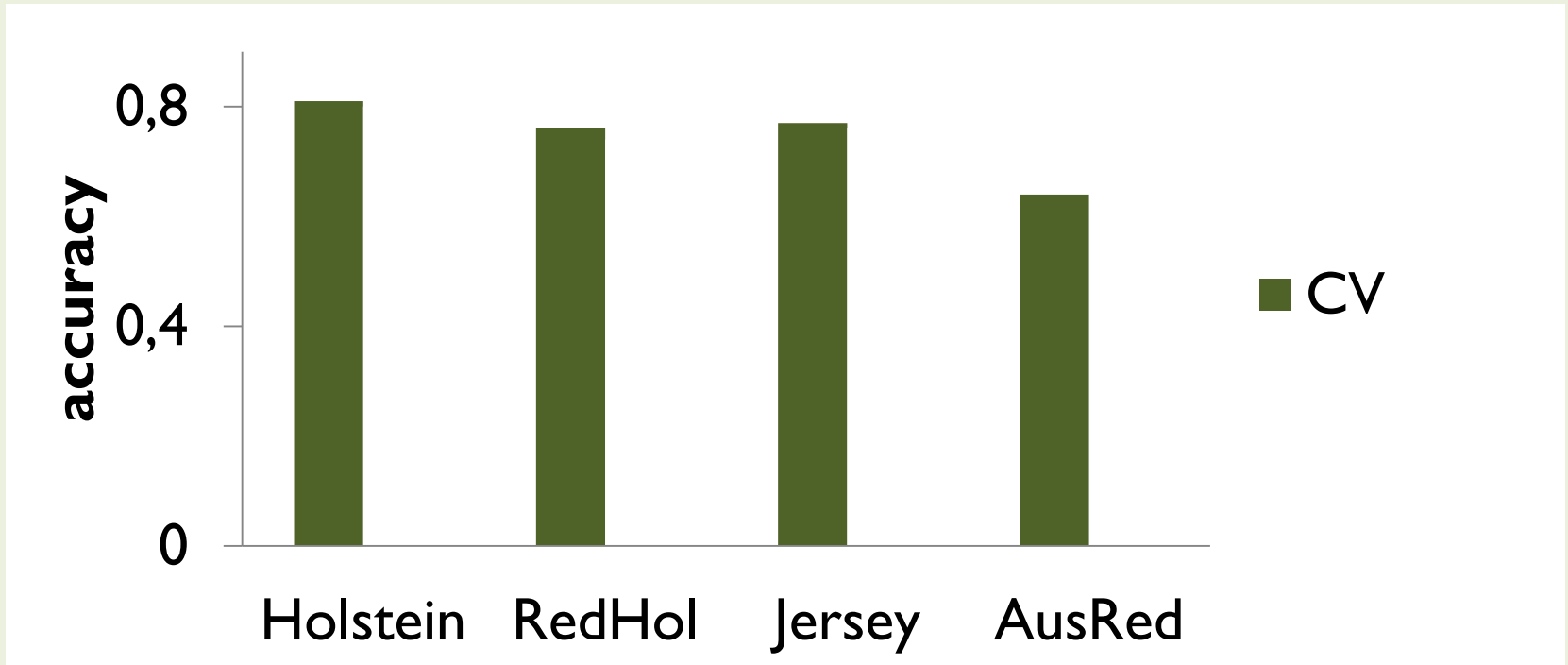# *Predicting the accuracy of the H+J reference population*

1) Estimate accuracies of single breed H and J reference populations by cross-validation, use those to estimate $\theta$,

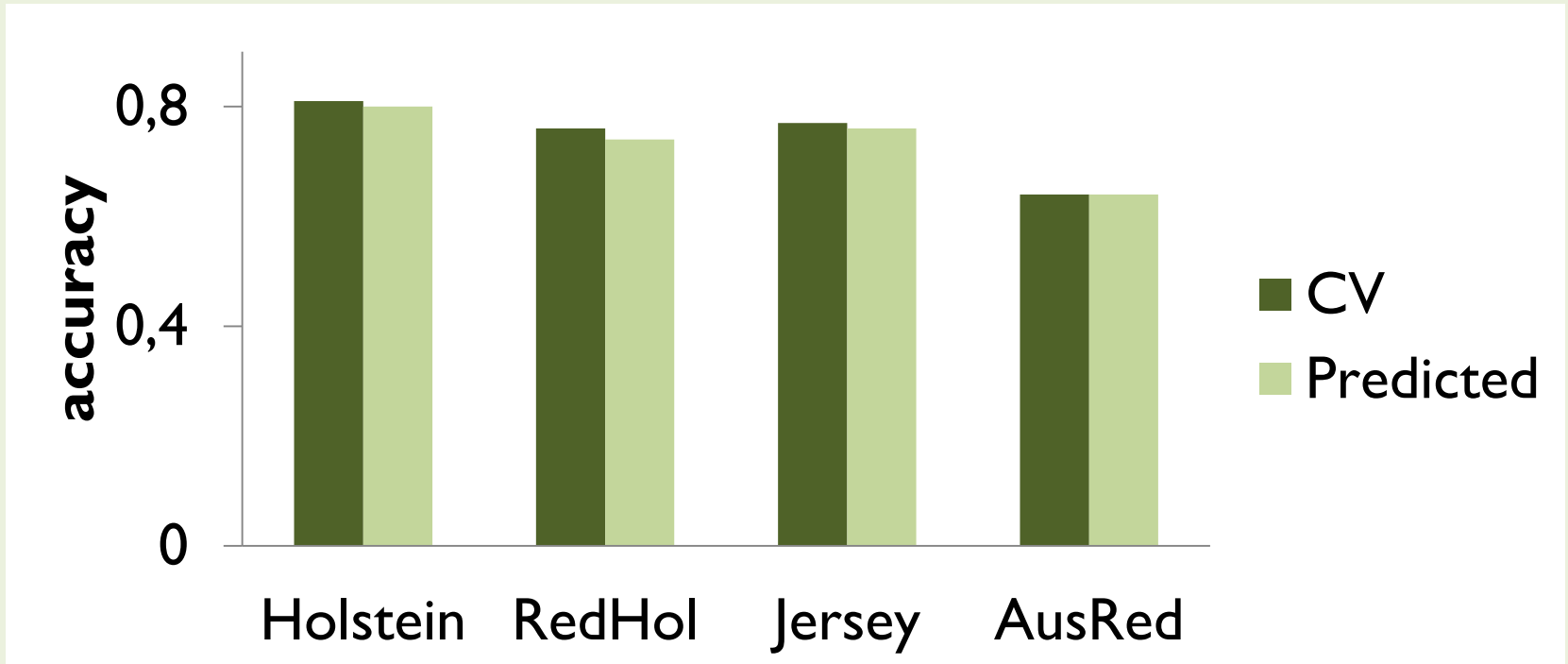$$\text{by solving } r = \sqrt{\frac{\theta}{1+\theta-r^2h^2}}$$

2) Add up single breed $\theta$'s to predict the accuracy of the multi breed reference population: $r_{H+J} =$

$$\sqrt{\frac{\theta_H+\theta_J}{1+\theta_H+\theta_J-r_{H+J}^2h^2}}$$

# Results
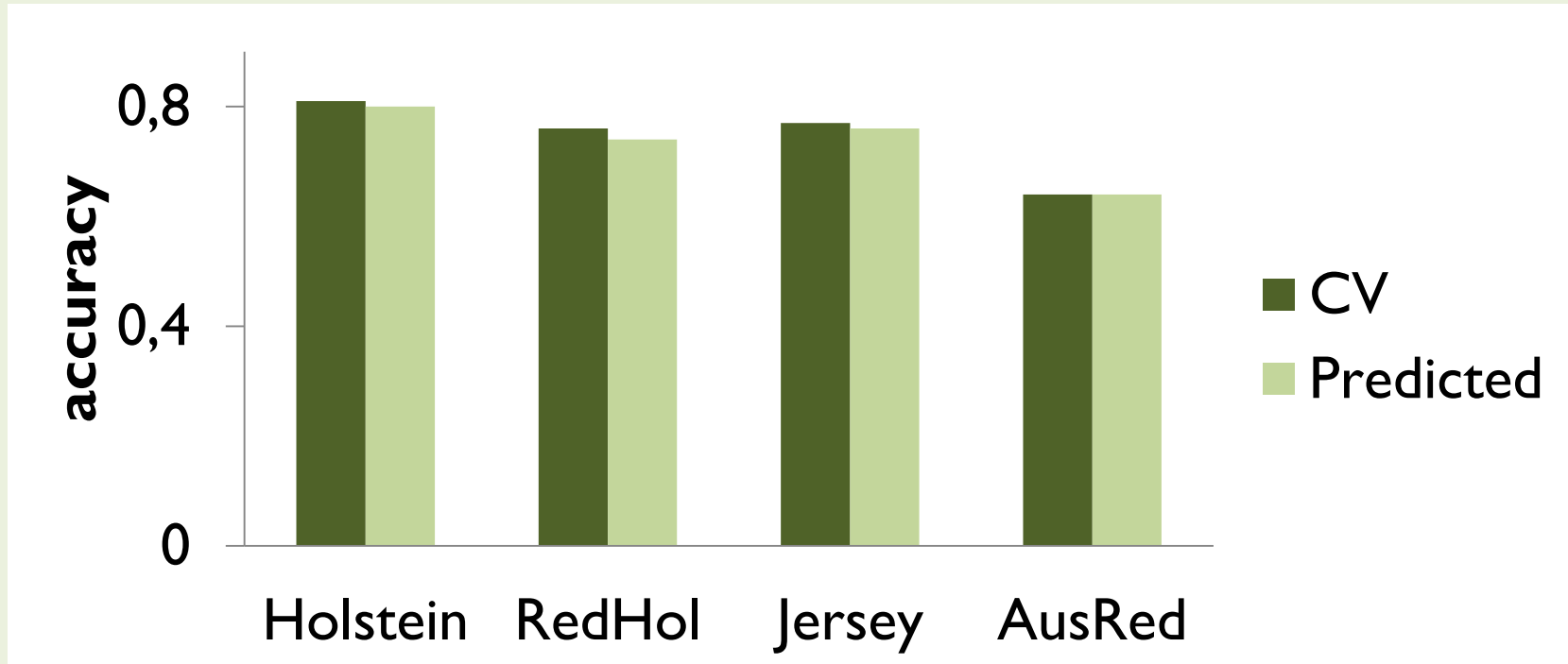
# Results

# Results



**Conclusion**  $\theta$'s obtained by cross-validation for single breed reference populations can be useful to predict the accuracy for a multi breed reference population